LA-UR-21-20546

Approved for public release; distribution is unlimited.

Title: Ergodic Theory and Dynamical Process Modeling

Author(s): Rupe, Adam Thomas

Intended for: Web

Issued: 2021-01-21

Ergodic Theory and Dynamical Process Modeling

Foundations for Continuum Computational Mechanics

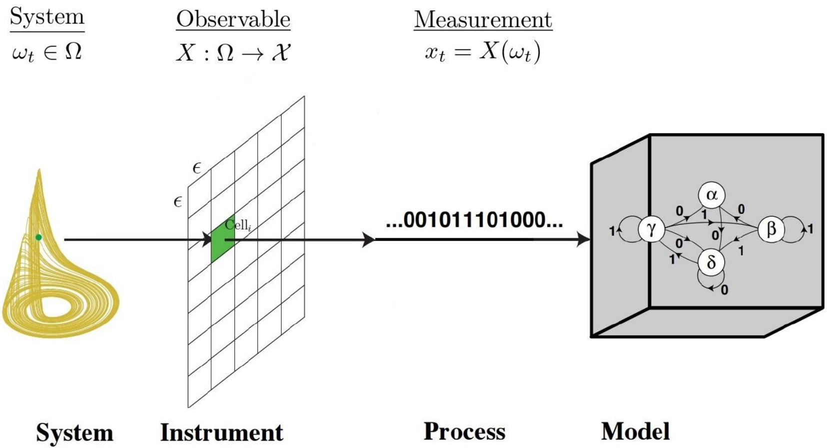
Adam Rupe

Center for Nonlinear Studies and Computational Earth Science
Los Alamos National Laboratory

Inference for Dynamical Systems Meeting, 2021

Nonlinear Modeling and Symbolic Processes

How does chaos generate randomness?



Systematic and rigorous method for converting a continuous dynamical system into a fully-discrete, i.e. *symbolic*, stochastic process.

Nonlinear Modeling and Symbolic Processes

Symbolic processes are the objects traditionally modeled by computational mechanics,

Studied through *ϵ-machines*

Unique minimal sufficient statistic of past for predicting the future, generated by *causal equivalence relation*:

$$\mathrm{past}_i \sim_\epsilon \mathrm{past}_j \iff \Pr(\mathrm{Future}|\mathrm{past}_i) = \Pr(\mathrm{Future}|\mathrm{past}_j)$$

- optimal prediction
- (causal) structure, organization
- directly calculate entropy rate
- process memory and complexity

Nonlinear Modeling and Symbolic Processes

When does nonlinear modeling work? – *generating partitions*

discrete-time dynamical system $(\Omega, \Phi : \Omega \to \Omega)$ – e.g. Poincare Map

$$\omega_{n+1} = \Phi(\omega_n)$$

partition phase space with *measurement function* $G_{\mathbb{P}} : \Omega \to \mathcal{A}$

$\mathbb{P}_i \cap \mathbb{P}_j = \emptyset$ and $\bigcup_{i=0}^{N} \mathbb{P}_i = \Omega$, and each partition carries unique symbol $a \in \mathcal{A}$

$\{\omega_0, \omega_1, \omega_2, \ldots\}$ becomes $\{a_0, a_1, a_2, \ldots\} = \{G_{\mathbb{P}}(\omega_0), G_{\mathbb{P}}(\Phi(\omega_0)), G_{\mathbb{P}}(\Phi^2(\omega_0)), \ldots\}$.

Nonlinear Modeling and Symbolic Processes

$G_{\mathbb{P}} \circ \Phi$ induces partition $\Phi^{-1}\mathbb{P}$ over Ω ; $(\Phi^{-1}\mathbb{P})_i$ is set of all $\omega \in \Omega$ s.t. $G_{\mathbb{P}}(\Phi(\omega)) \in \mathbb{P}_i$

each time step induces new $\Phi^{-n}\mathbb{P}$ whose elements are $\omega \in \Omega$ s.t. $G_{\mathbb{P}}(\Phi^n(\omega)) \in \mathbb{P}_i$

partition refinement $\mathbb{P} \vee \mathbb{Q} = \{\mathbb{P}_i \cap \mathbb{Q}_j \ : \ \mathbb{P}_i \in \mathbb{P} \text{ and } \mathbb{Q}_i \in \mathbb{Q}\}$ also a partition

first *dynamical refinement* of $\mathbb{P}$ under Φ is $\mathbb{P} \vee \Phi^{-1}\mathbb{P}$,
maps point $\omega \in \Omega$ to two-symbols $a_0 a_1 \in \mathcal{A} \times \mathcal{A}$

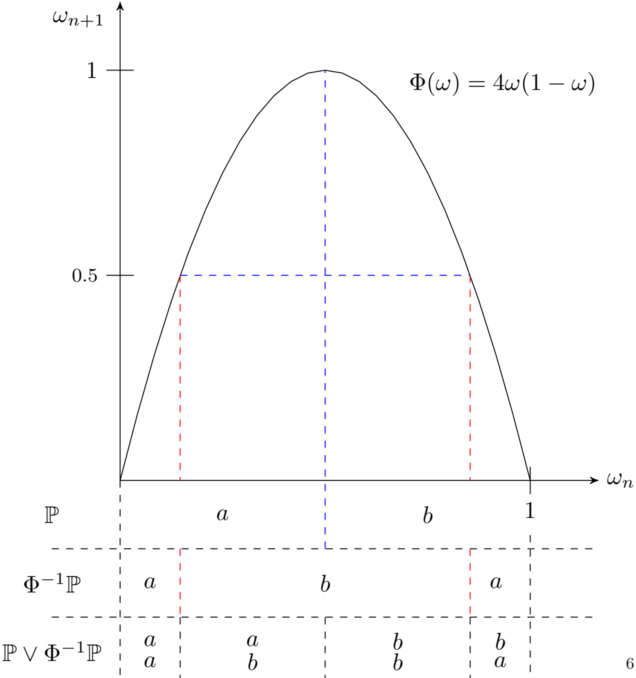
the *full dynamical refinement* of $\mathbb{P}$ under Φ is $\mathbb{P} \vee \Phi^{-1}\mathbb{P} \vee \Phi^{-2}\mathbb{P} \vee \Phi^{-3}\mathbb{P} \cdots$,
maps point $\omega \in \Omega$ to infinite-length symbol sequence $\mathcal{A} \times \mathcal{A} \times \mathcal{A} \times \cdots$

a *generating partition* is a partition $\mathbb{P}$ s.t. the full dynamical refinement is a.e. one-to-one between points $\omega \in \Omega$ and infinite-length symbol sequences—volume of partition elements goes to 0 for full dynamical refinement

A.N. Kolmogorov, Russian Academy of Sciences (1959), Y.G. Sinai Russian Academy of Sciences (1959)

Generating Partition of Logistic Map

$$G_{\mathbb{P}}(x) = \begin{cases} a & 0 \le \omega \le 0.5 \\ b & 0.5 \le \omega \le 1 \end{cases}$$



Kolmogorov-Sinai Entropy: Randomness from Chaos

entropy of partition (in terms of invariant distribution over partition elements)

$$H(\mathbb{P}) = -\sum_i \Pr(\mathbb{P}_i) \log \Pr(\mathbb{P}_i)$$

entropy rate

$$h_\nu(\Phi, \mathbb{P}) = \lim_{N \to \infty} \frac{1}{N} H\left(\bigvee_{n=0}^{N} \Phi^{-n} \mathbb{P} \right)$$

Kolmogorov-Sinai (metric) entropy

$$h_\nu(\Phi) = \sup_{\mathbb{P}} h_\nu(\Phi, \mathbb{P})$$

achieved for generating partitions—provides variational principle for approximation

(asymptotic distribution over $\mathbb{P}_i$ limits to invariant density f^* of Φ for generating $\mathbb{P}$)

Pesin's theorem: relation to positive Lyapunov exponents

$$h_\nu(\Phi) \leq \sum_{\lambda_i > 0} \lambda_i$$

Y. Pesin. (1977) Russian Math Surveys

Nonlinear Modeling and Symbolic Processes

Symbolic process is a *shift dynamical system* $(\mathcal{A}^{\mathbb{Z}}, \sigma)$ under *shift operator*

for two symbol sequence $x, y \in \mathcal{A}^{\mathbb{Z}}$, $y = \sigma(x) \iff y_i = x_{i+1}$ for all $i \in \mathbb{Z}$,
i.e. σ advances observation time index forward

$$
\begin{tikzcd}
\underline{\text{System}} & \underline{\text{Measurement Symbol}} & \underline{\text{Symbol Sequence}} \\
\omega_n \in \Omega \arrow[r, "G_{\mathbb{P}} : \Omega \to \mathcal{A}"] \arrow[d, "\Phi"] & x_0 \in \mathcal{A} & \cdots x_{-2} x_{-1} x_0 x_1 x_2 \cdots \arrow[l, "\mathbf{0}"'] \arrow[d, "\sigma"'] \\
\omega_{n+1} = \Phi(\omega_n) \arrow[r, "G_{\mathbb{P}} : \Omega \to \mathcal{A}"] & y_0 = x_1 \in \mathcal{A} & \cdots y_{-2} y_{-1} y_0 y_1 y_2 \cdots \arrow[l, "\mathbf{0}"']
\end{tikzcd}
$$

Converted nonlinear dynamical system into (symbolic) measurement process governed by **linear*, infinite-dimensional operator* σ that advances observation time

D. Lind and B. Marcus (1995) Cambridge University Press

Systems, Data, and Models

Advantages of nonlinear modeling:

- symbolic processes allow for discrete information and computation theory
- clean and rigorous framework for information storage, generation, and processing
- interpretability of system structure and organization through ϵ-machines

Challenges:

- even for idealized systems, generating partitions hard to find (e.g. Henon map)
- for a given physical system, don't have full control of system measurements
- scalability of inference and interpretability for large alphabets

generalize to continuum setting using Koopman and Perron-Frobenius Operators

A. Lasota and M. C. Mackey (2013). Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics. Springer

System: $\omega_0 \in \Omega$

Observations: $x_0 = X(\omega_0)$

$X : \Omega \to \mathcal{X}$

$X_t : \Omega \to \mathcal{X} = U^t X$

$= X \circ \Phi^t$

Φ^t

U^t

$X : \Omega \to \mathcal{X}$

Credit: NASA–Apollo 17 crew

Systems and Data

The "true" physical system described by dynamical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$:

- phase space Ω is complete metric space (typically $\mathbb{R}^d$ or d-dimensional manifold)
- Σ_Ω a σ-algebra (Borel sets)
- ν a reference measure (Borel or Lebesgue)—phase space volume
- Φ the generator of (semi)group of measurable flow maps $\{\Phi^t : \Omega \to \Omega\}$

$$\Phi = \lim_{\tau \to 0} \tfrac{1}{\tau}\big(\Phi^t(\omega) - \Phi^{t+\tau}(\omega)\big)$$

- orbits $\{\omega(t) \; : \; t \in \mathbb{R}_{(\geq 0)}\}$ continuous in time
- discrete intervals are bounded $||\Phi^{t+\delta}(\omega_0) - \Phi^t(\omega_0)|| < \epsilon$

Systems and Data

The "observed" or "measured" system is measurable space $(\mathcal{X}, \Sigma_{\mathcal{X}})$

Observable $x \in \mathcal{X}$ generated by the dynamical system under the measurable mapping $X : \Omega \to \mathcal{X}$ so that $x_t = X(\omega_t)$

Generally interested in *partially-observable systems*, for which X is *not invertible*: knowledge of x insufficient for determining state ω of the true system
$\implies$ there are "unobservable" or "immeasureable" degrees of freedom in ω

Will later consider a second set of observables $(\mathcal{Y}, \Sigma_{\mathcal{Y}})$ that are also given as a (generally non-invertible) measurable map $Y : \Omega \to \mathcal{Y}$ s.t. $y_t = Y(\omega_t)$.

X (and Y) represent "windows" through which we can view true physical system, but can never have a full view with $\mathcal{X} = \Omega$

Asymptotic behavior of ω may be reconstructable from x: *delay-coordinate embedding*

N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw (1980) F. Takens (1981)

Koopman Operators

system "observable" $f : \Omega \to \mathbb{C}$, element of a function space (typically $L^\infty(\Omega, \nu)$ or $L^2(\Omega, \nu)$)

Koopman operators $\{U^t : \mathcal{F} \to \mathcal{F}\}$ evolve observables through composition with Φ^t

$$U^t f = f \circ \Phi^t$$
$$f_t(\omega) \equiv f\big(\Phi^t(\omega)\big) = U^t f(\omega)$$

linear, infinite-dimensional operators whose action on observable $f \in \mathcal{F}$ gives the time shifted observable (function) $f_t = U^t f$

Inherits (semi)group structure of $\{\Phi^t\}$ s.t. $U^t \circ U^{\Delta t} = U^{t+\Delta t}$ and generated by

$$U f = \lim_{t \to 0} \frac{1}{t}\left(U^t f - f\right)$$

X (and Y) vector-valued observables, i.e. $X_i = f$, evolved by product operator U^t

Dynamical Processes

observables may be collected in a *time series* $\{x_0, x_1, \ldots, x_{T-1}\}$—a time-ordered sequences of "measurements" of x taken at uniform time intervals.

a *dynamical process* is bi-infinite time series of observables $\{\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots\}$

* for non-invertible dynamics, index is "observation time"

in terms of Koopman operators:

$$x_t = X(\omega_t) = X\big(\Phi^t(\omega_0)\big) = X_t(\omega_0) = U^t X(\omega_0)$$

Goal of dynamical process modeling: infer or approximate action of U on observables X

$$\{x_0, x_1, x_2, x_3, \ldots\} = \{U^0 X(\omega_0), U^1 X(\omega_0), U^2 X(\omega_0), U^3 X(\omega_0)\}$$

(finite) *reconstructability*: X in finite invariant subspace $\mathbf{U}$ s.t. $U^t X \in \mathbf{U}$ for all t

S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. N. Kutz (2016). PloS One

Perron-Frobenius Operators

Rather than evolve observables, *Perron-Frobenius operators* $P^t : L^1(\Omega, \nu) \to L^1(\Omega, \nu)$ evolve *densities* $f \in L^1(\Omega, \nu)$: $f \geq 0$ and $||f|| = 1$.

$$f_t = P^t f$$

$$\int_S P^t f d\nu = \int_{\left(\Phi^t\right)^{-1}(S)} f d\nu \qquad \text{for } S \in \Sigma_\Omega$$

May also consider P^t evolving $L^2(\Omega, \nu)$ measures μ: $\mu_t = P^t \mu$

relation to densities through Radon-Nikodym:

$$\mu^f(S) = \int_S f d\nu \quad \text{and} \quad f = \frac{d\mu^f}{d\nu}$$

$P^t : L^1 \to L^1$ adjoint of $U^t : L^\infty \to L^\infty$ and $P^t : L^2 \to L^2$ adjoint of $U^t : L^2 \to L^2$

$$\langle P^t f, g \rangle = \langle f, U^t g \rangle$$

Stochastic Processes

Can now formulate *continuous stochastic processes* generated by dynamical systems

- dynamical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$ with initial probability density (ensemble) f_0
- density at time t is $f_t = P^t f_0 \implies$ probability measure $\mu_t(S) = \int_S f_t d\nu$
- at time t have probability space $(\Omega, \Sigma_\Omega, \mu_t) \implies$ observable map $X : \omega_t \mapsto x_t$ now defines random variable X_t
- X_t distributed according via pushforward $\mu_t^X(S_\mathcal{X}) = \mu_t(X^{-1}(S_\mathcal{X}))$ for $S_\mathcal{X} \in \Sigma_\mathcal{X}$

$$\mathrm{pr}(X_t \in S_\mathcal{X}) = \int_{S_\mathcal{X}} d\mu_t^X = \int_{X^{-1}(S_\mathcal{X})} d\mu_t = \int_{X^{-1}(S_\mathcal{X})} f_t \, d\nu = \int_{X^{-1}(S_\mathcal{X})} P^t f_0 \, d\nu$$

Therefore, an initial density f_0 on a dynamical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$ produces continuous stochastic process $\{X_0, X_1, X_2, \dots\}$

random variables are actually $X(t, \omega)$; for fixed ω a *sample path* given by $t \mapsto X(t, \omega)$, here are (continuous) dynamical processes $\{x_0, x_1, x_2, \dots\}$

Connection to Symbolic Processes: Koopman

symbolic process:

$$\{\omega_0, \omega_1, \omega_2, \ldots\} \rightarrow \{a_0, a_1, a_2, \ldots\} = \{G_{\mathbb{P}}(\omega_0), G_{\mathbb{P}}(\Phi(\omega_0)), G_{\mathbb{P}}(\Phi^2(\omega_0)), \ldots\}$$

dynamical process:

$$\{\omega_0, \omega_1, \omega_2, \ldots\} \rightarrow \{x_0, x_1, x_2, \ldots\} = \{U^0 X(\omega_0), U^1 X(\omega_0), U^2 X(\omega_0), \ldots\}$$

Logistic map:

Partition isomorphic to generating partition given by

$$G_{\mathbb{P}}(x) = \begin{cases} 0 & 0 \leq x \leq 0.5 \\ 1 & 0.5 \leq x \leq 1 \end{cases}$$

$G_{\mathbb{P}}(x)$ generally given as sum of labeled indicator functions for partition elements $\mathbb{P}_i$:

$$G_{\mathbb{P}}(x) = \sum_{i=0}^{N-1} i \mathbb{1}_{\mathbb{P}_i}, \quad N = ||\mathbb{P}||$$

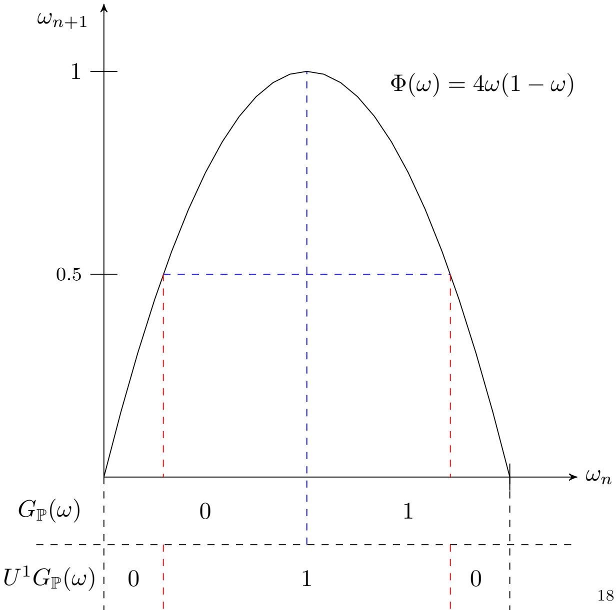
$$\mathbb{1}_{\mathbb{P}_i} = \begin{cases} 1 & x \in \mathbb{P}_i \\ 0 & x \notin \mathbb{P}_i \end{cases}$$

Connection to Symbolic Processes: Koopman

$$G_{\mathbb{P}}(\omega) = \begin{cases} 0 & 0 \leq \omega \leq 0.5 \\ 1 & 0.5 \leq \omega \leq 1 \end{cases}$$

$$U^1 G_{\mathbb{P}}(\omega) = G_{\mathbb{P}}\big(\Phi(\omega)\big) = \Phi^{-1}\mathbb{P}$$

$$U^1 G_{\mathbb{P}}(\omega) = \begin{cases} 1 & \frac{1-\sqrt{\frac{1}{2}}}{2} \leq \omega \leq \frac{1+\sqrt{\frac{1}{2}}}{2} \\ 0 & \text{otherwise} \end{cases}$$



Connection to Symbolic Processes: Perron-Frobenius

reference measure vs invariant measure

if reference measure is invariant, constant density $f = 1$ is stationary: $P^t 1 = 1$

for logistic map with Borel / Lebesgue reference measure:

$$P^1 1 = \frac{1}{2\sqrt{1-\omega}}$$

$\implies$ Borel measure not invariant

invariant measure μ^* found from solving $P^1 f^* = f^*$:

$$\mu^*(S) = \int_S f^* d\omega = \int_S \frac{d\omega}{\pi\sqrt{\omega(1-\omega)}}$$

Ulam and von Neumann (1947)

Connection to Symbolic Processes: Perron-Frobenius

reference measure vs invariant measure

consider uniform density on $\mathbb{P}_1$ $\big(\Pr(a_0 = 1) = 1\big)$ and its evolved density

$$f(\omega) = \begin{cases} 0 & 0 \le \omega \le 0.5 \\ 2 & 0.5 \le \omega \le 1 \end{cases} \qquad\qquad P^1 f(\omega) = \frac{1}{2\sqrt{1-\omega}}$$

with Borel reference measure:

$$\Pr(a_1 = 1) = \quad 2\int_{\frac{1}{2}}^{\frac{1+\sqrt{\frac{1}{2}}}{2}} d\omega = \int_{\frac{1}{2}}^{1} \frac{d\omega}{2\sqrt{1-\omega}} = \sqrt{\frac{1}{2}}$$

with invariant reference measure:

$$\Pr(a_1 = 1) = \quad 2\int_{\frac{1}{2}}^{\frac{1+\sqrt{\frac{1}{2}}}{2}} d\mu^* = 2\int_{\frac{1}{2}}^{\frac{1+\sqrt{\frac{1}{2}}}{2}} \frac{d\omega}{\pi\sqrt{\omega(1-\omega)}} = \frac{1}{2}$$

Ergodicity

Typical to consider *measure-preserving* dynamical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$ with $\Phi\nu = \nu$

Assumes a *stationary* stochastic process: $\mathrm{pr}(X_t \in S_\mathcal{X}) = \mathrm{pr}(X_0 \in S_\mathcal{X})$ for all t

This corresponds to dynamics on an *attractor*, but want to consider more general dynamics that include transient relaxation to the attractor

We do this using *ergodic components*, which correspond to *basins of attraction* (including the attractor itself)

A dynamical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$ is *ergodic* if $\nu(S) = 0$ or $\nu(S) = 1$ for every *invariant set* $S : \Phi^{-1}(S) = S$
$\implies$ all invariant sets are trivial subsets of Ω and we must study Φ on the entire space Ω

*** Note ergodicity is independent of measure preservation—but they are often assumed together

Attractors and Basins

invariant set : $\Phi^{-1}(S) = S$
forward-invariant set : $\omega_0 \in S \implies \Phi^t(\omega_0) \in S$ for all t

an invariant set is necessarily forward-invariant, but converse not true

An *attractor* of $(\Omega, \Sigma_\Omega, \nu, \Phi)$ is a set $A \in \Sigma_\Omega$ with

- A is a forward-invariant set of Ω under Φ
- there exists an open set $B \supset A$, call the *basin of attraction* of A, s.t. for every $\omega \in B$ $\lim_{t \to \infty} \Phi^t(\omega) \in A$, and
- there is no proper subset of A with the first two properties

Attractors are *not* Φ-invariant, but *basins are Φ-invariant*

Can define basins as limit of pre-images of attractor: $B = \lim_{t \to \infty} \left(\Phi^t\right)^{-1} A$

Attractor Basins as Ergodic Components

Multi-stable dynamical systems with multiple attractors can be partitioned: each basin may be treated independently since, by definition, orbits never cross basin boundaries

Without loss of generality, will from here out consider systems that either:

- have one attractor with Ω as basin
- independently consider *ergodic components*: reduced systems with $\Omega = B$

System $(\Omega, \Sigma_\Omega, \nu, \Phi)$ considered this way is thus *ergodic*

Asymptotic behavior on the attractor:

- ergodicity guarantees P^t has unique invariant density $f^* : P^t f^* = f^*$
- this defines *asymptotic invariant measure* $\mu^*(S) = \int_S f^* d\nu$
- ergodic theorem: time averages equal phase space averages

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(\Phi^k(\omega)\right) = \frac{1}{\mu^*(\Omega)} \int_\Omega f(\omega) d\mu^*$$

The Prediction Problem

We set up a general prediction problem for dynamical processes as follows:

For a physical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$ let the observable $X : \Omega \to \mathcal{X}$ represent the collection of all possible measurements that can be made of the system.

Let $Y : \Omega \to \mathcal{Y}$ represent "variables of interest"

* typically $\mathcal{Y} \subseteq \mathcal{X}$, but variables of interest may not be adequately measurable, in which case $\mathcal{Y} \nsubseteq \mathcal{X}$
* variables in $\mathcal{X}$ that are not in $\mathcal{Y}$ are sometimes called *exogeneous*

From definitions / assumptions, the physical system $(\Omega, \Sigma_\Omega, \nu, \Phi)$ generates the *true* dynamical processes (indices again correspond to observation time)

$$\{\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots\}$$
$$\{\ldots, y_{-2}, y_{-1}, y_0, y_1, y_2, \ldots\}$$

The Prediction Problem

Denote *past* random variable $\overleftarrow{X} = \{\ldots, X_{-2}, X_{-1}, X_0\}$ with realizations $\overleftarrow{x} = \{\ldots, x_{-2}, x_{-1}, x_0\}$ represent observed series of measurements up to present time t_0

Similarly denote *future* random variable $\overrightarrow{Y}_\tau = \{Y_1, Y_2, Y_3, \ldots, y_\tau\}$ with realizations $\overrightarrow{y}_\tau = \{y_1, y_2, y_3, \ldots, y_\tau\}$ represent future values of variables of interest to lead time τ

deterministic prediction problem: find *target function* $\mathcal{T}_\tau$ that maps $\overleftarrow{x}$ to $\overrightarrow{y}_\tau$

$$\text{minimize} \;\; ||\mathcal{T}_\tau \circ \overleftarrow{x} - U^\tau Y||_{L^2(\nu)}$$

R. Alexander and D. Giannakis, Physica D: Nonlinear Phenomena (2020)

probabilistic prediction problem: find *conditional distribution* $\mathrm{pr}(\overrightarrow{Y}_\tau | \overleftarrow{X} = \overleftarrow{x})$

Predictive Models of Dynamical Processes

Model dichotomy:

myopic models—learn target functions for finite τ (and finite-length pasts)

- (N)ARIMA(X)
- Analogue forecasting
- Neural networks (e.g. LSTM, TCN, RC, etc.)

process models—generative models that approximate U (i.e. $\lim \tau \to \infty$)

- approximate action of U with surrogate dynamical system $\widetilde{\Phi}_\alpha : \mathcal{X} \to \mathcal{X}$
- numerical simulations with discretization X and approximation scheme α
- complex physics simulations with analysis/assimilation $\widetilde{X}$ from measurements X with parameterizations α
- reduced-order models (ROMS)
- (generalized) Galerkin approximations of U

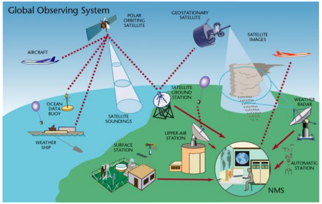
Example: Earth System

(analysis / assimilation)

System

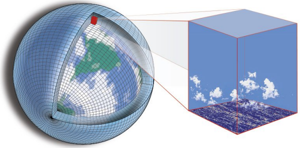
$\xrightarrow{X : \Omega \to \mathcal{X}}$

Observations



$\xrightarrow{\widetilde{X}}$

Simulation



“data image”

Φ^t

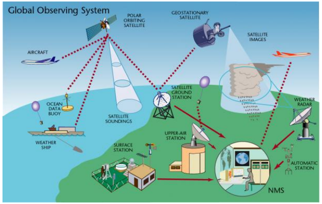
$X_t = U^t X$

U^t

$\widetilde{\Phi}^t_\alpha$

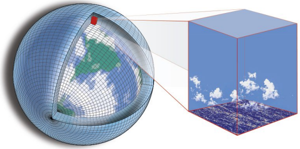
Credit: NASA–Apollo 17 crew

$\xrightarrow{X : \Omega \to \mathcal{X}}$



Credit: WMO

$\xrightarrow{\widetilde{X}}$



Credit: Tapio Schneider/Kyle Pressel/Momme Hell/Caltech

Galerkin Approximations

EDMD: project U^t onto finite subspace spanned by dictionary $\mathbf{\Psi} = [\psi^1, \ldots, \psi^k]^{\mathrm{T}}$

$$
\begin{tikzcd}
\underline{\text{System}} & \underline{\text{Observations}} & \underline{\text{Galerkin Approximation}} \\
\omega_0 \in \Omega \arrow[r, "X : \Omega \to \mathcal{X}"] \arrow[dd, "\Phi^t"] \arrow[ddr, "X_t = U^t X"] & x_0 = X(\omega_0) \arrow[r, "\mathbf{\Psi}"] \arrow[dd, "U^t"'] & {[\psi^1(x_0), \ldots, \psi^k(x_0)]^{\mathrm{T}}} \arrow[d, "\mathbf{U}^t"'] \\
 & & {[\mathbf{U}^t\psi^1(x_0), \ldots, \mathbf{U}^t\psi^k(x_0)]^{\mathrm{T}}} \arrow[d, dashed, leftrightarrow, "\text{minimize}"] \\
\omega_t = \Phi^t(\omega_0) \arrow[r, "X : \Omega \to \mathcal{X}"] & x_t = X(\omega_t) \arrow[r, "\mathbf{\Psi}"] & {[\psi^1(x_t), \ldots, \psi^k(x_t)]^{\mathrm{T}}}
\end{tikzcd}
$$

M. O. Williams, I. G. Kevrekidis, and C. W. Rowley (2015). Journal of Nonlinear Science
S. Klus, P. Koltai, and C. Schütte (2016). Journal of Computational Dynamics